

基于表示学习的车辆到达时间预测

李昕彤[†], 李茂源, 陈茜雅, 司浩田

(东北林业大学 奥林学院, 黑龙江 哈尔滨 150006)

摘要:随着智慧交通的兴起,人们对高效率出行的需求日益增加,因而寻求更优秀的模型来估计车辆行驶时间成为首要任务。由于交通系统有较强的非线性,并且受天气、时间等多种因素影响,综合考虑历史车流量信息和当前路况,提出了一种基于表示学习的多模态拟合模型。将车辆行驶时间估计(ETA)问题视为一个基于一组含有大量有效特征的纯时空序列的回归问题,分别采用不同的机器学习模型来解决每一部分的回归问题。通过滴滴出行的数据来训练模型,充分利用 SDNE (Structure Deep Network Embedding)、LSTM (Long Short-Term Memory)、xDeepFM (eXtreme Deep Factorization Machine) 算法的各自优势,最后的对比实验表明,提出的模型优于传统的深度学习算法。

关键词:智慧交通;表示学习;多模态拟合;机器学习

中图分类号:TP391.9

文献标识码:A

Prediction of Vehicle Reaching Time Based on Representation Learning

LI Xintong[†], LI Maoyuan, CHEN Xiya, SI Haotian

(Aulin College, Northeast Forestry University, Harbin, Heilongjiang 150006, China)

Abstract: With the rise of smart transportation, people's demand for efficient travel is increasing. Therefore, the search for better models to estimate vehicle travel time has become a top priority. Since the traffic system has a strong nonlinearity and is affected by various factors such as weather and time, this paper proposes a multimodal fitting model based on representation learning, considering historical traffic flow information and current road conditions. This article regards the estimated time of arrival (ETA) problem as a regression problem based on a set of pure spatiotemporal sequences with a large number of effective features, and uses different machine learning models to solve each part of the regression problem. The model is trained through the data of Didi, and the respective advantages of SDNE (Structure Deep Network Embedding), LSTM (Long Short-Term Memory), xDeepFM (eXtreme Deep Factorization Machine) algorithms are fully utilized. Finally, the comparison test shows that the model proposed in this article is superior to the traditional deep learning algorithm.

Key words: smart transportation; representation learning; multi-modal fitting; machine learning

近年来,智慧交通的发展正在改变、影响着人们的出行方式,人们开始习惯了出门时软件中提示的车辆到达时间和路径规划,滴滴出行、高德地图等众多出行软件帮助人们更高效地到达目的地,节

约出行时间、降低出行成本,随着用户数量的增多和用户需求增加,精确处理庞大的车辆行程时间估计(ETA)请求成了研发热点^[1]。精确处理 ETA 请求是出行软件为用户提供高质量出行服务的保

障之一。

针对此问题,越来越多的学者和科学工作者开始关注道路拥堵的解决方案,投入了前沿的技术并针对 ETA 预测算法提出了种种优化,获得了许多成果。2016 年,谷歌首先提出并研究开发了 Wide & Deep 模型^[2],兼具了逻辑回归和深度神经网络的优点,为后来推荐系统的很多模型奠定了基础。2018 年,Wang 等^[3]最早提出了基于深度学习的车辆到达时间预测算法 DeepTTE (Deep Temporal Trajectory Embedding),采用多任务学习的方式训练,使用卷积神经网络捕获轨迹的位置信息,同时将语义信息编码后与各个路段位置信息拼接,通过 LSTM 神经网络预测各个路段所需时间及总时间,通过调整两者比例以获得较好的预测效果。2018 年,Zhang 等^[4]提出 Deep Travel 算法,首次采用网格图的方式编码轨迹,将轨迹分解为网格序列并拼接编码后的其他语义信息,依次输入 BiLSTM (Bidirectional Long Short-Term Memory) 获取轨迹的表示向量,避免了轨迹序列过长导致的信息丢失。Wang 等^[5]提出 WDR (Wide Dynamic Range) 算法,首次采用推荐系统算法 WD (Wide & Deep Learning) 与 LSTM 相结合的方式预测车辆到达目的地所需时间。Wide 部分、Deep 模型、RNN (Recurrent Neural Network) 分别提取出对应的轨迹信息,拼接后使用全连接网络输出到达目的地所需时间;该算法已被滴滴出行用于预测出租车到达目的地所需时间。2020 年,Shen 等^[6]提出 TTPNet (Travel Time Predict Network) 算法。将所有轨迹序列嵌入网格图中,利用张量分解算法恢复各个网格 1 h 内的速度信息,获取网络节点表示信息后拼接其他语义信息并输入 BiLSTM 网络,得到预测结果。2019 年,Li 等^[7]提出为不同功能区域定制不同的预测模型。2020 年,Li 等^[8]提出了一种新的时空融合图神经网络(STFGNN)用于交通流预测,有效地学习了隐藏关系。2021 年,Hui 等^[9]提出使用空间注意力机制来自适应地捕获不同路段之间的动态相关性。2022 年,Ji 等^[10]首次提出使用神经网络配合物理约束和复杂动力学来进行交通流量预测。2022 年,Yang 等^[11]考虑到研究对象之间互动的影 响,提出了一种基于时空图的卷积网络算法进行有效预测。

上述这些文献通过对算法本身进行优化提升最终的预测准确度,但是由于道路信息、车辆情况复杂且实时变化,仅仅靠优化算法是不够的,利用传统的方法和单一的框架来预测结果依旧不够理

想,因此本文分别为路网信息模块、车辆轨迹模块、其他离散语义信息模块和预测模块选用不同的算法来提高准确度,提出了一种基于表示学习的机器学习算法模型,用以预测车辆沿某路线到达目的地所需时间。

1 模型设计

1.1 特征提取

滴滴出行等应用程序的广泛使用使得每天都会产生大量的车流量数据,较为复杂、冗余的原始信息难以直接输入神经网络,且不加筛选地输入信息会使训练结果较差,因此为了保证模型的准确性和有效性,笔者首先为基于位置的系统建立高维特征映射。并把出行过程中的特征信息归纳为以下几类:空间信息,时间信息、交通信息、其他离散语义信息。

空间信息。出行所需的时间长短与驾驶车辆途经城市的主要路线长度和其地理区域大小密切相关。因此,需要抽象实际中的城市主要道路信息,提取出路网信息序列的空间特征集。城市路网信息一般由多条城市道路、路口信息和实时交通信号数据组成,可以通过进一步分类将其抽象化为有向带权图,笔者首先将出行的车辆轨迹信息映射到基础道路网络结构中,得到各路段和各交叉口的序列。然后提取出构成路网的所有构件的特征信息,如主要道路、十字路口和各类交通信号灯等等。此外,笔者还进一步提取计算了特定区域内道路的相关信息点,以便更准确的表示路网空间信息,更好地对出行时间进行预测。

时间信息。出行选择的时间段往往是直接影响车辆总体旅程时间的另一个关键因素;例如,同样一条路线,高峰期出行的时间花费就比其余时间段要多;商业区附近的路段,周末时的出行时间会大于平常工作日时间段。因此,笔者以几种不同的时间特征值来表示该旅程中的实际出发时间,包括每天的时间段、假日指标和高峰时间指标等。

交通信息。道路网络中的交通状况直接影响出行路线的选择和出行时间。因此,笔者利用 ACM SIGSPATIAL 2021 GISCU 比赛中的实时交通动态系统数据,每 5 min 即可为实验提供交通路网模型中对每个路段的实时交通速度估计。然后,笔者在特征图中分别为每个划分的路段使用这几种类型的旅行速度,如实时估计的速度、平均速度和自由流速度等。

其他离散语义信息。除了上面列举的信息外,还存在一些语义信息会影响出行路线的选择和出

行时间,例如:个性化信息、轨迹总长度、出发时间、天气状况、气温。不同的司机可能对路线选择有不同的偏好,且开车习惯会有较大差别,因此,笔者在其他离散语义信息中引入个性化信息,将每一个用户与其驾驶车辆的车牌号相对应,通过车牌号来记忆每一位用户的驾驶偏好,以便根据其偏好推荐不同的行驶路线。此外,对于天气状况,笔者将其分为 1—5 个级别,同时记录星期数、最高气温和最低气温。所要求提取对象的连续特征一般包括实数值、离散值特征等多种形式组合的特征。最后,整理分析了收集到的特征信息,发现最终共获得了数百个类别和数百万个维度中的一组特征。

1.2 构建模型

基于提取的特征,笔者将 ETA 学习问题重新改写成一个更为标准化的机器学习形式。首先分别用 $y = [y_1, y_2, \dots, y_N]$ 来表示每个出行样本的旅行时间的基本时间真值标签,用 $y_i = e_i - s_i \in R_+$ 计算每次出行样本的总行程时间,计算的基本方式为每次到达最终目的地时间 e_i 和下次出发旅行的基本时间 s_i 之间对应的时间间隔。分别用 $X = [x_1, x_2, \dots, x_N]$ 来表示样本,其中每个样本 $x_i \in R^d$,将道路网络中沿路线路径的轨迹解码为 d 维向量。因此,概括来说,实验的目标就是训练一个模型,它可以准确地预测未来未知数据 $x_q \notin X$ 的出行时间。

统计滴滴打车已有的用户订单信息能够发现,用户好评率和满意度实际都与模型估计出的平均时间差以及用户计划出行总时间有关,因此,平均绝对百分比误差(MAPE)是一个相对更精确、更合理可靠的度量标准。实验的主要目标是将 MAPE 最小化

$$\min_f \sum_{i=1}^N \frac{|y_i - f(x_i)|}{y_i} \quad (1)$$

其中, $f(x_i)$ 为路径 x_i 的 ETA,函数 f 为回归模型。但是,上述两种形式之间的 MAPE 可能存在发生过拟合事件的潜在风险,因此,为了尽量保证模型对未知数据具有更良好的可预测分析性能,笔者尝试通过适当引入某些额外的正则化项来避免过拟合的发生,一旦发生过拟合,需要立刻停止训练,修改模型,从而约束模型的复杂性。修改后,笔者将 ETA 问题的一般优化目标表示为

$$\min_f \sum_{i=1}^N \frac{|y_i - f(x_i)|}{y_i} + \Omega(f) \quad (2)$$

其中, $\Omega(f)$ 是控制模型复杂性及防止过高拟合性的正则化项。这里使用梯度增强学习决策树 GBDT (Gradient Boosted Decision Trees) 来进行学习。在 GBDT 中,预测模型可以表示为一个加

性树模型,即

$$f(x) = \sum_{i=1}^T f_i(x) \quad (3)$$

因此,给定一个输入样本 x ,树模型返回 x 所属的叶子的分数,其中 T 表示树的数量。GBDT 模型 f 的复杂性受树的结构控制。

该模型的优化问题是一个凸优化问题。然而,由于 MAPE 函数是不可微的,所以目标是不可光滑的。这里利用次梯度法解决优化问题。

1.3 联合模型

在大量实际工程应用过程中,笔者发现 GBDT、SVR (Support Vector Regression)、普通神经网络和随机森林均不是解决 ETA 问题的最优选择。GBDT 很难适应大型特征集,随机森林容易出现过拟合现象且计算成本较高、训练时间较长。另一方面,笔者有大量的数据分布复杂的历史流量数据,单一的模型难以很好地适配所有数据,因此,笔者采用深度学习技术来适应这个问题。

该实验的模型联合训练 SDNE 模型、LSTM 神经网络和 xDeepFm 模型,以结合不同模型各自的优点。该模型结构如图 1 所示。

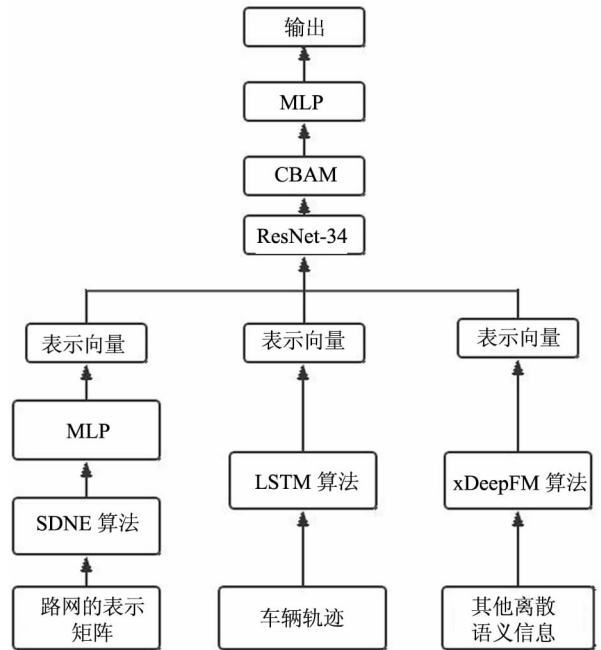


图 1 联合模型

实验模型首先考虑将输入的特征信息投影映射到高维的特征空间中。这一步也可以通过计算两个输入特征之间的交叉乘积来实现,这就类似于 FM 中的特征交互作用。变换后,可以直接应用仿射变换 $y = w \cdot x + b$ 。

深度模型中首先需要利用特征嵌入层来将稀疏的输入特征转换为密集输入特征。这种嵌入是

通过定义一个密集的特征向量(大小为 20)来表示每一类高维稀疏特征。然后,将这些密集的输出特征与嵌入的输出特征连接起来,并直接馈入一个前馈神经网络单元 FNN (Feed Forward Neural Network)阵列中,也通常被称为多层感知器(Multi Layer Perceptron, MLP)阵列。顶层模型的回归器结合了宽模型和深度模型的输出来帮助提供最终结果的预测。通常用稀疏特征表示分类信息,用密集特征表示数值信息。

上述 3 个模型要求最后获得的向量尺寸相同,这样才可以按通道(channel)拼接起来,用 1 维残差 ResNet-34 (Residual Network-34)做卷积操作,最后拼接 CBAM 注意力模块(Convolutional Block Attention Module),并用 MLP 输出预测值。

1.4 路网的图表示

城市路网一般由城市道路、路口和交通信号灯组成,可以进一步将其功能抽象为有向带权图,其中每个节点对应一个表示向量,整个路网对应一个表示矩阵,由各个节点表示向量拼接所组成。

训练路网图表示时笔者采用 SDNE 算法、Graph2Vec (Graph to Vector)算法等进行实验,并最终选择效果最好的 SDNE 算法。输入路网表示矩阵,并借助 SDNE 算法,最终用 MLP 全连接神经网络输出一个表示向量。

1.5 车辆轨迹的表示学习

每个轨迹由一系列节点序列组成,各节点的信息由 1 个向量组成,向量内容包括:节点长度、5 min 内节点内车辆平均速度以及第一步得到的节点表示向量(图表示向量),最终通过 LSTM 算法输出一个表示向量。

LSTM 的逻辑核心就是细胞状态,在实现 LSTM 逻辑的每个推理步骤过程中,输入门、忘记门、输出门和调制输入都被更新如下:

$$i_t = \sigma(W_i[x_t; h_{t-1}] + b_i) \quad (4)$$

$$f_t = \sigma(W_f[x_t; h_{t-1}] + b_f) \quad (5)$$

$$o_t = \sigma(W_o[x_t; h_{t-1}] + b_o) \quad (6)$$

$$g_t = \tanh(W_g[x_t; h_{t-1}] + b_g) \quad (7)$$

最后,将内存单元格和隐藏状态每次更新即可。

1.6 其他离散语义信息的表示学习

除去路网信息和车流量信息外,还有一些其他因素会影响车辆到达时间,包括时间、天气以及司机个人喜好等等,这里将这些所有因素归结为离散语义信息。每条语义信息包括轨迹总长度、出发时间(每 5 min 一个序号,一天共 288 个序号)、天气状况(分为 1-5 个等级)、星期数、最高气温、最低气温、车辆 ID(由于每个司机驾驶 1 辆车,因此可

以将司机和车辆捆绑起来,即用车辆 ID 记录司机的驾驶习惯和行驶偏好)。这里采用 xDeepFM 算法来进行训练。

xDeepFM 是 Wide & Deep 的改进版,在此基础上添加了压缩感知层(Compressed Sensing Network, CIN)显式的构造有限阶特征组合。总体结构中有 3 个分支:Linear 层(输入稀疏的 01 向量)、DNN (Deep Neural Network)层(输入经过 embedding 的稠密向量)和 CIN 层。其中 CIN 层是关键,其输入所有 field 列向量横向拼接得到的矩阵,每次输入的第 i 层矩阵都会和原始输入做交互,得到 $i+1$ 层的特征矩阵。由于每一层只包含某一阶的特征组合,所以每一层的特征矩阵都会用于对最终结果作预测,将最终每层得到的所有标量拼接得到 1 维向量,作为 CIN 层的输出。然后再与 Linear 和 DNN 部分的输出拼接,得到最终的预测结果,输出一个表示向量。计算公式如下:

$$X_{h,*}^k = \sum_{i=1}^{H_{k-1}} \sum_{j=1}^m W_{ij}^{k,h} (X_{i,*}^{k-1} \circ X_{j,*}^0) \quad (8)$$

1.7 时间预测模块

前面 3 个步骤依次对每个模块选择适合的模型算法进行了预测,最终得到 3 个同尺寸的表示向量,将这 3 个向量按 channel 拼接,通过 1 维 ResNet 做卷积操作,并最终拼接 CBAM 注意力模块,通过 MLP 输出预测值,其中采用的损失函数是平均绝对误差百分比(MAPE)。

ResNet 残差卷积网络可以很好地解决随着网络的加深,训练难度加大而产生的梯度消失现象,以及由于越深的网络返回的梯度相关性会越差,接近白噪声,而导致梯度更新也接近随机扰动的问题。其引入了一个深度残差框架,即让卷积网络节点去学习残差的映射,而不是仅仅期望每一个堆叠层上的网络节点都能很完美地学习去拟合一个潜在的映射。

必须假设模型中包含的所有其他重要参数都是能在没有 MAPE 损失的条件下直接使用反向传播进行联合训练所获得的。由于 3 个模块的结合,使得实验难以找到一个相对较为合适的全局学习率。因此,笔者选择了具有自适应步长和动量的随机梯度下降方法(Adaptive Moment Estimation, Adam)对模型进行优化。

2 数据集

实验采用了 ACM SIGSPATIAL 2021 GIS-CUP 比赛数据,数据内容为 2021 年 8 月及 9 月 1 日深圳市出租车轨迹数据。其中 8 月数据用作

训练集,9月1日数据用作测试集。当司机响应顾客的请求时,收集一个样本,直到司机接到顾客。当乘客上车时,开始收集出行样本,直到乘客到达目的地。从真实场景中收集到的车辆轨迹可能极其复杂,包含多种类型的线路。准确无误地预测这条路线的最佳旅行时间是一个具有挑战性且意义深远的问题。在去除旅行时间异常的情况后,实验获得了最终的原始数据。由于大多数 ETA 模型存在数据稀疏性问题,经过分析给定数据集上的数据覆盖率的统计数据,得出结论,滴滴司机的轨迹在白天覆盖了约 45% 的道路连接,覆盖范围在黎明前降至最低水平。

3 实验结果

训练得到的结果表明 MAMPE 值为 0.10,低于同期进行处理的其他模型,其他模型的训练结果在对比实验中进行展示。

4 对比实验

为了证明本模型的优良效果,笔者同步进行对比实验。运用同样的数据,通过传统的深度学习算法进行预测,并将预测结果和实验结果做对比,对比结果如表 1 所示:

表 1 模型对比	
模型	MAPE
GBDT	0.49293
SVR	0.50127
普通神经网络	0.381349
随机森林	0.162037
本文模型	0.10

从表 1 可以清楚地看出这几种模型中本文所设计的联合模型的 Mape 值最低,实验效果最好,即该算法框架可以更有效、更精确地解决 ETA 问题。

5 结 论

通过研究道路交通网和车流量信息,建立了基于表示学习的算法框架,实现了分模块分别学习的多模态处理,不同模块数据选择不同的算法而达到

总体性能最优。对比实验均表明,新的算法模型具有最小的 MAPE 值,这充分说明本算法相较于以往传统的算法,在模型质量方面有了较大进步,该新模型在快速解决 ETA 学习计算问题方面比笔者现有的深度学习计算模型更有效,实现了更精确的出行时间预测,节省了用户时间,同时在某种程度上缓解了一部分交通压力。虽然这种新技术的学习模型相较于现如今普遍使用的模型有了一定的进步,但是该模型也依旧存在可以改进的巨大空间,笔者相信,通过进一步丰富学习特征系统模型和引入一些性能更强大的学习系统模型,能有效地提高 ETA 学习系统的准确性,为用户的出行提供更好的体验。

参考文献

[1] 陈凌源,高上进,谢振宇. 基于神经网络的到港 ETA 预测模型[J]. 航空学报,2013,34(4):873—880.

[2] CHENG H T,KOC L,HARMSSEN J,et al. Wide & deep learning for recommender systems[J]. ACM,2016,59(1):7—10.

[3] WANG D,ZHANG J B,CAO W ,et al. When will you arrive? estimating travel time based on deep neural networks [C]. AAAI Conference on Artificial Intelligence,2018,32(1):3654—3660.

[4] ZHANG H Y,WU H,SUN W W,et al. Deeptravel: a neural network based travel time estimation model with auxiliary supervision. [J]. IEEE Transactions on Intelligent Transportation Systems,2018,20(1):365—374.

[5] WANG Z,FU K,YE J P,et al. Learning to estimate the travel time [J]. IEEE Transactions on Intelligent Transportation Systems,2018,20(3):858—866.

[6] SHEN Y B,JIN C Q,HUA J X. TTPNet: a neural network for travel time prediction based on tensor decomposition and graph embedding[J]. IEEE Transactions on Knowledge and Data Engineering,2020,(99):1—1.

[7] LI Y X,ZHENG Y,YANG Q. Efficient and effective express via contextual cooperative reinforcement learning. [J]. SIGKDD explorations,2019,25(3330968):510—519.

[8] LI M Z,ZHU Z X. Spatial-temporal fusion graph neural networks for traffic flow forecasting[J]. AAAI Conference on Artificial Intelligence,2020,35(5):4189—4196.

[9] HUI B,YAN D,CHEN H Q,et al. TrajNet: a trajectory-based deep learning model for traffic prediction[A]. Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining [C]. 2021,12(6):716—724.

[10] JI J H,WANG J Y,JIANG Z,et al. STDEN: towards physics-guided neural networks for traffic flow prediction[J]. ArXiv,2022,36(4):4048—4056.

[11] YANG J K,HAN C. Pedestrian trajectory prediction based on improved social spatio-temporal graph convolution neural network[J]. ACM Transactions on Intelligent Systems and Technology,2022,13(3):63—67.